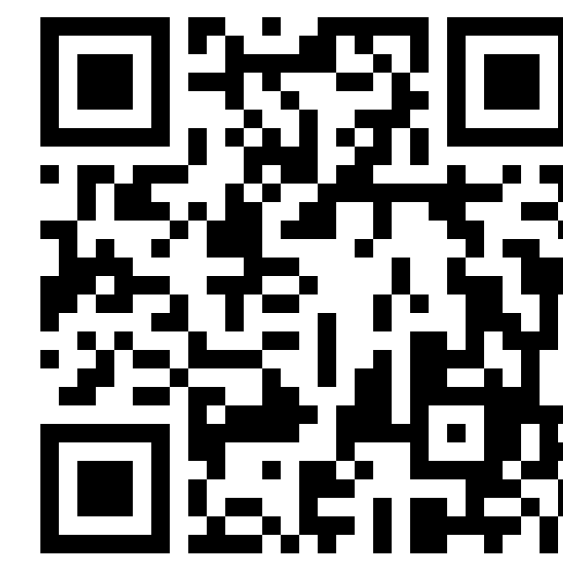


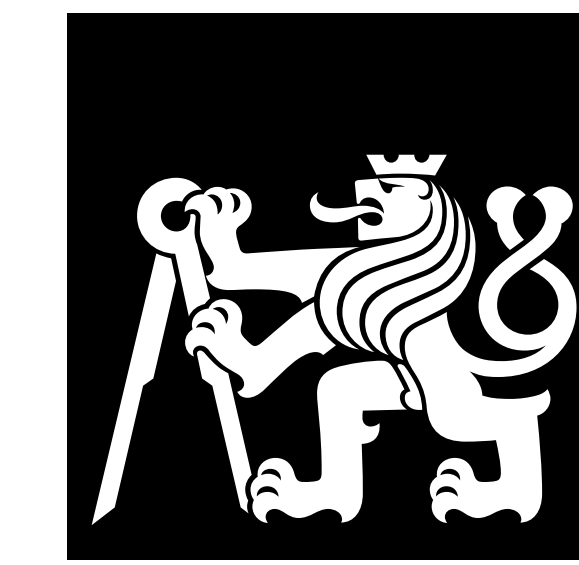
Study On Using Language Models In Storytelling Games

Author
Ing. Jiří Macháček

Supervisor
Ing. Jan Matoušek



**HaLLMark
Prototype**



**FACULTY
OF INFORMATION
TECHNOLOGY
CTU IN PRAGUE**

Motivation

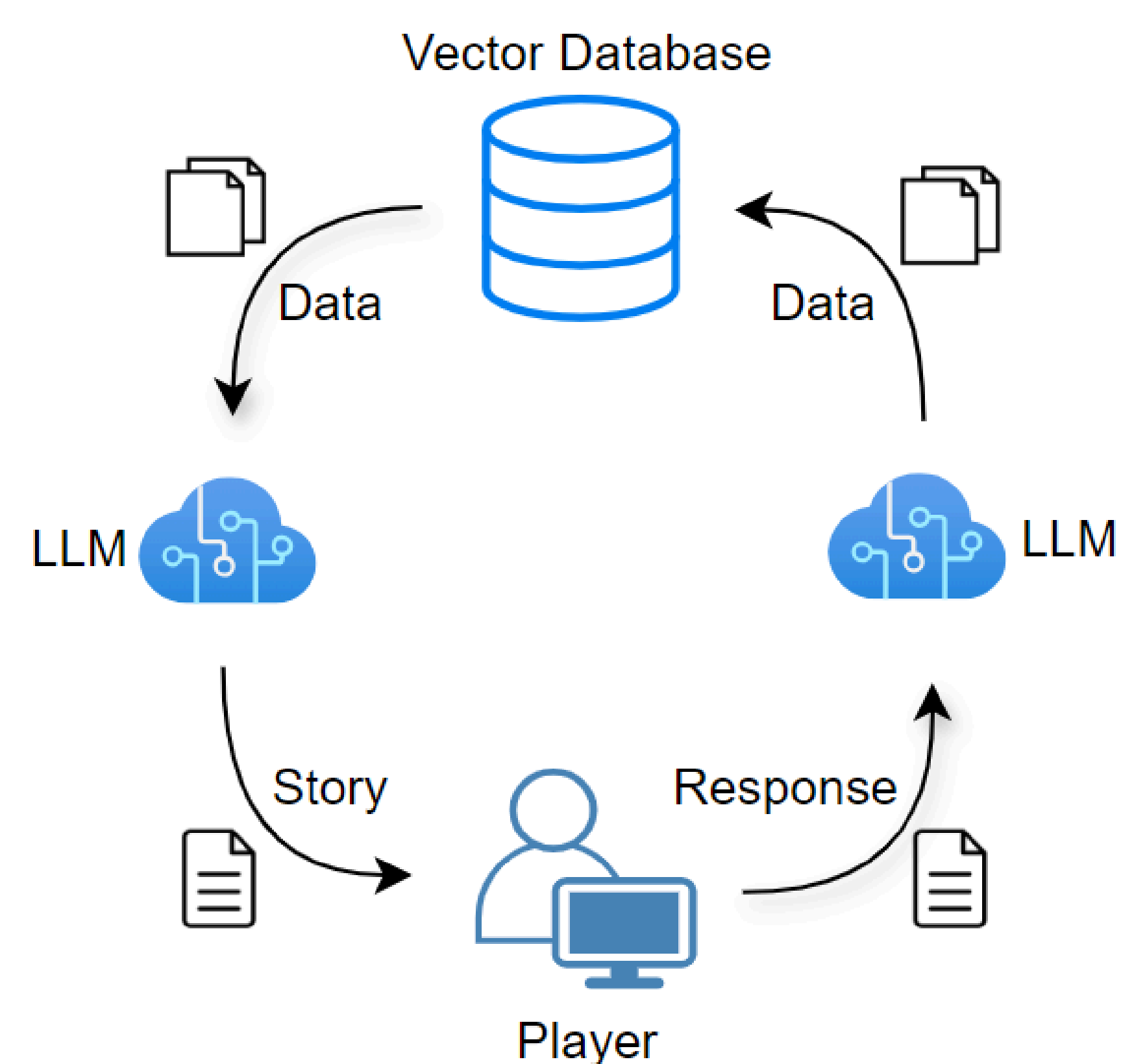
In story-based games, the interactions between the player and other game characters **are usually limited** to only a few options the game designers prepared for the players. With such a low number of possibilities, the players can quickly **get frustrated** because none of the available options suit them, and they would instead do something else entirely.

Large language models (LLMs) represent a powerful tool for text generation (among others), which can be used to **create stories**. Large language models can also **impersonate various characters** and react adequately to anything the user tells them.

We aim to use large language models to tell game stories as game characters **with which the player can interact freely**.

Creating Entertaining Stories

Every exciting story must consist of at least one **conflict**, the **story tension must gradually increase** and the story's world and characters must remain **consistent**. Remaining consistent is most challenging when the storyteller is a large language model, and to overcome this shortcoming, we suggest using **vector databases**.



Whenever an essential fact about a character or the game world is presented, it gets saved into a vector database. When a conflict or character's speech is being generated, **the most important facts are queried** from the vector database to provide **more context** for the generator.

Generated stories are further improved by:

- ◆ Using **prompt engineering** techniques
- ◆ Mentioning characters already existing in the story
- ◆ Including **plot twists** in the story

HaLLMark the Storytelling Prototype

You see, my cart, full of the most exquisite silks, was recently overtaken by a group of mischievous Dreamland gnomes.

Problem presented by the king's subject

Oh, Your Majesty, those gnomes are a tricky bunch! Perhaps offer them a deal: a silk hat for each gnome, in exchange for returning the rest of the silk.

Jester's advice on how to deal with the problem

Offer the gnomes to keep 20 silk hats in exchange for the rest of the goods. If they refuse, send our royal army to deal with the gnomes.

Player's decision on how to deal with the problem

The gnomes gleefully accepted the offer of silk hats, but a surprising turn of events unfolded when they used the hats to create colorful hot-air balloons, soaring away with the remaining silk goods.

The army, unprepared for aerial adversaries, stood helpless as the gnomes' laughter echoed through Dreamland.

Consequences of the player's decision

Using the most modern knowledge about LLMs and the most popular storytelling techniques, the game HaLLMark has been implemented.

As the king of a fantasy kingdom, the player's goal is to solve the issues of their subjects. Using **textual input**, the king can freely discuss the problems with his **subject or advisor**. After deciding how to deal with the presented issue, the consequences of the player's decisions are generated, and all the **vital information is saved** to be used further in the game. LLMs are also responsible for many other game mechanics.

The prototype **gained positive feedback** during player testing, with testers mainly highlighting the previously unseen game design concept, **story consistency and originality**. This thesis shows that it is possible to **tell certain game stories** and **impersonate game characters** using LLMs and modern techniques related to them, like vector databases and prompt engineering.